Exploratory Data Analysis (EDA) — Iris Dataset
Tools: Python, pandas, matplotlib

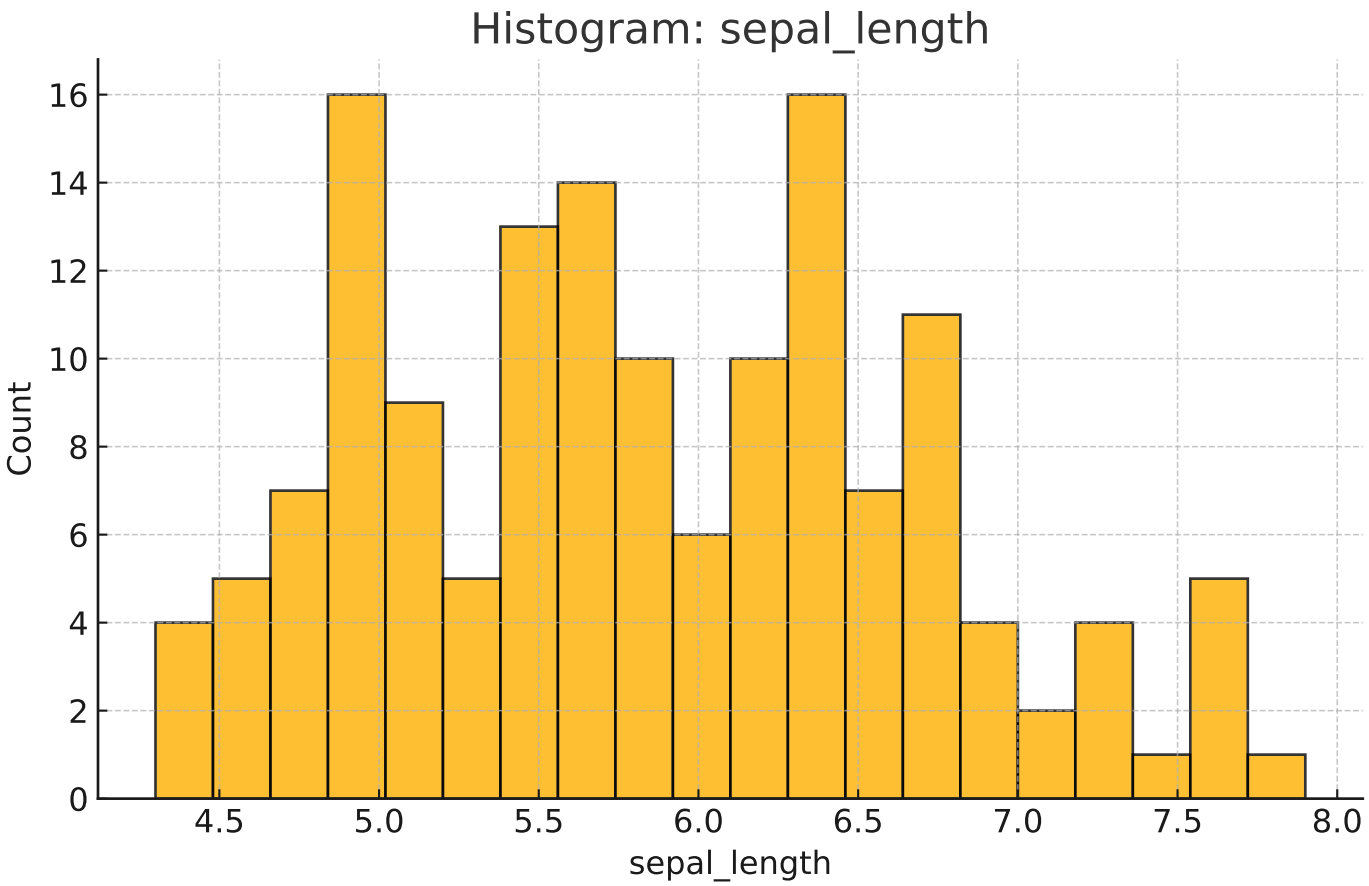
Dataset overview:
• Samples: 150
• Features: 4 numeric + 1 categorical ('species')
• Classes: setosa, versicolor, virginica

Descriptive Statistics (.describe())

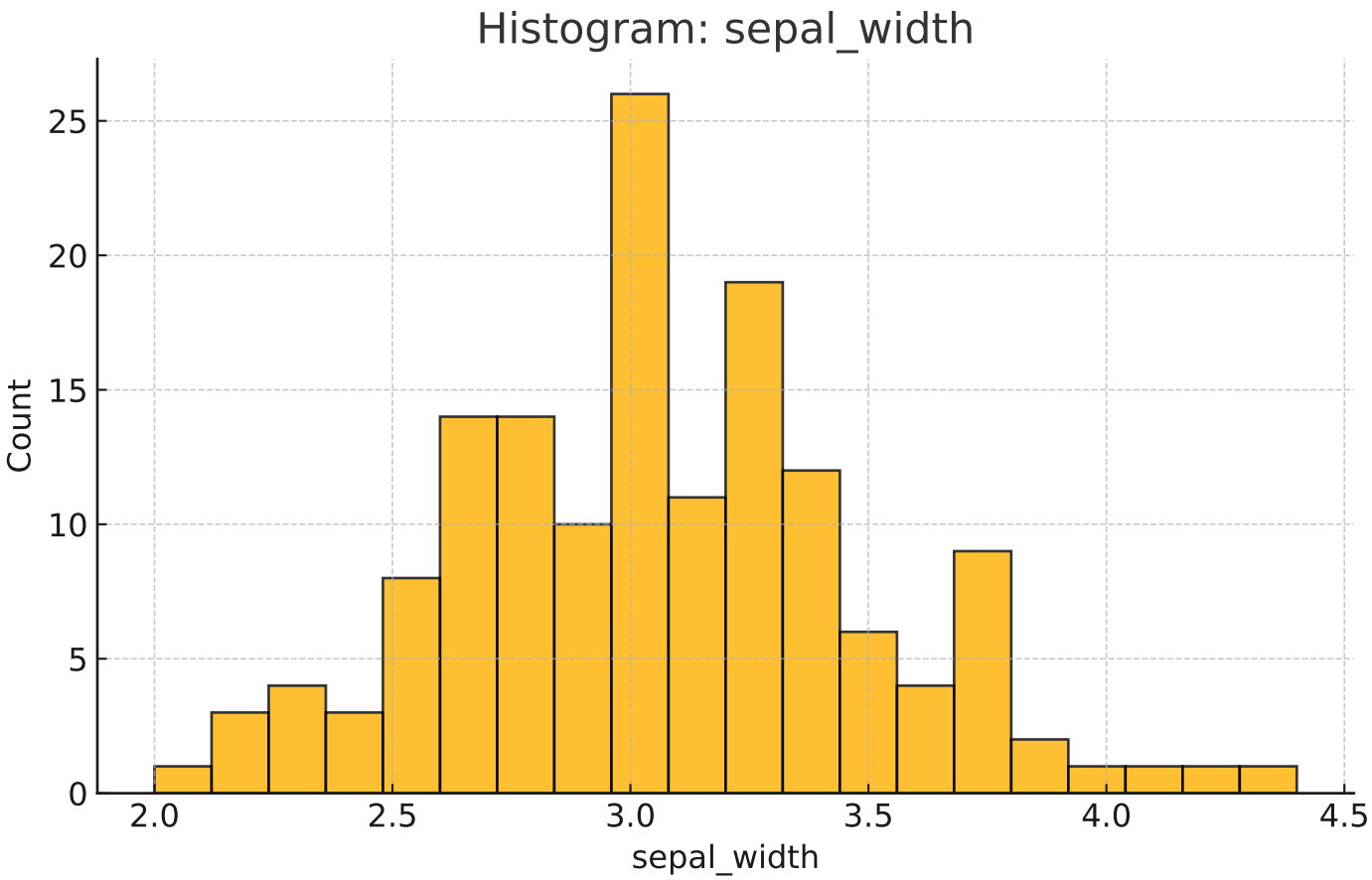
	sepal_length	sepal_width	petal_length	petal_width	species
count	150.000	150.000	150.000	150.000	150
unique	NaN	NaN	NaN	NaN	3
top	NaN	NaN	NaN	NaN	setosa
freq	NaN	NaN	NaN	NaN	50
mean	5.843	3.057	3.758	1.199	NaN
std	0.828	0.436	1.765	0.762	NaN
min	4.300	2.000	1.000	0.100	NaN
25%	5.100	2.800	1.600	0.300	NaN
50%	5.800	3.000	4.350	1.300	NaN
75%	6.400	3.300	5.100	1.800	NaN
max	7.900	4.400	6.900	2.500	NaN

DataFrame Info (.info())

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   sepal_length  150 non-null    float64
 1   sepal_width   150 non-null    float64
 2   petal_length  150 non-null    float64
 3   petal_width   150 non-null    float64
 4   species       150 non-null    object 
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
Class Balance (.value_counts())
setosa        50
versicolor    50
virginica     50
```

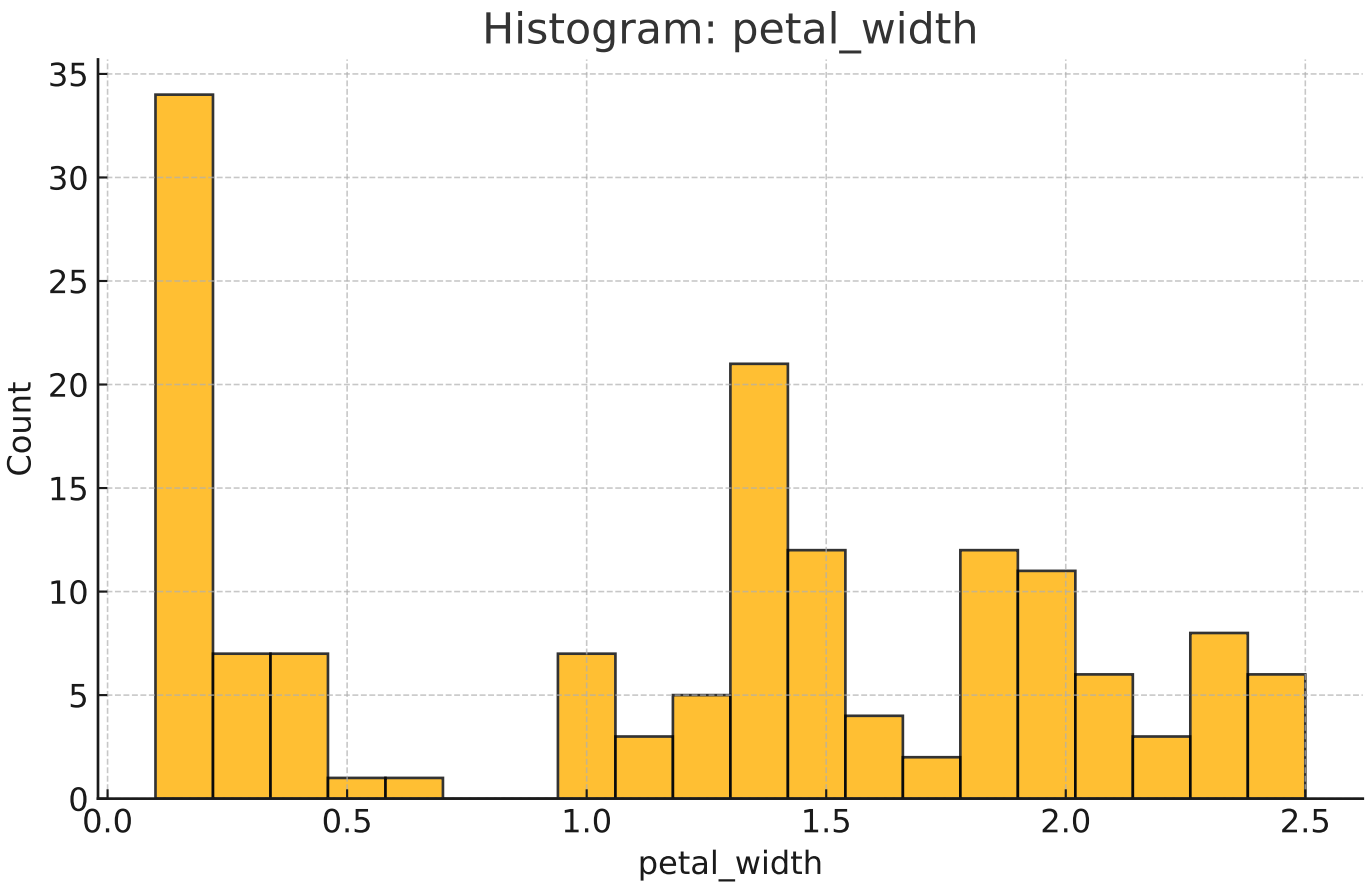


Observations:
Sepal length is roughly bell-shaped with most values between ~5.0 and 6.5.

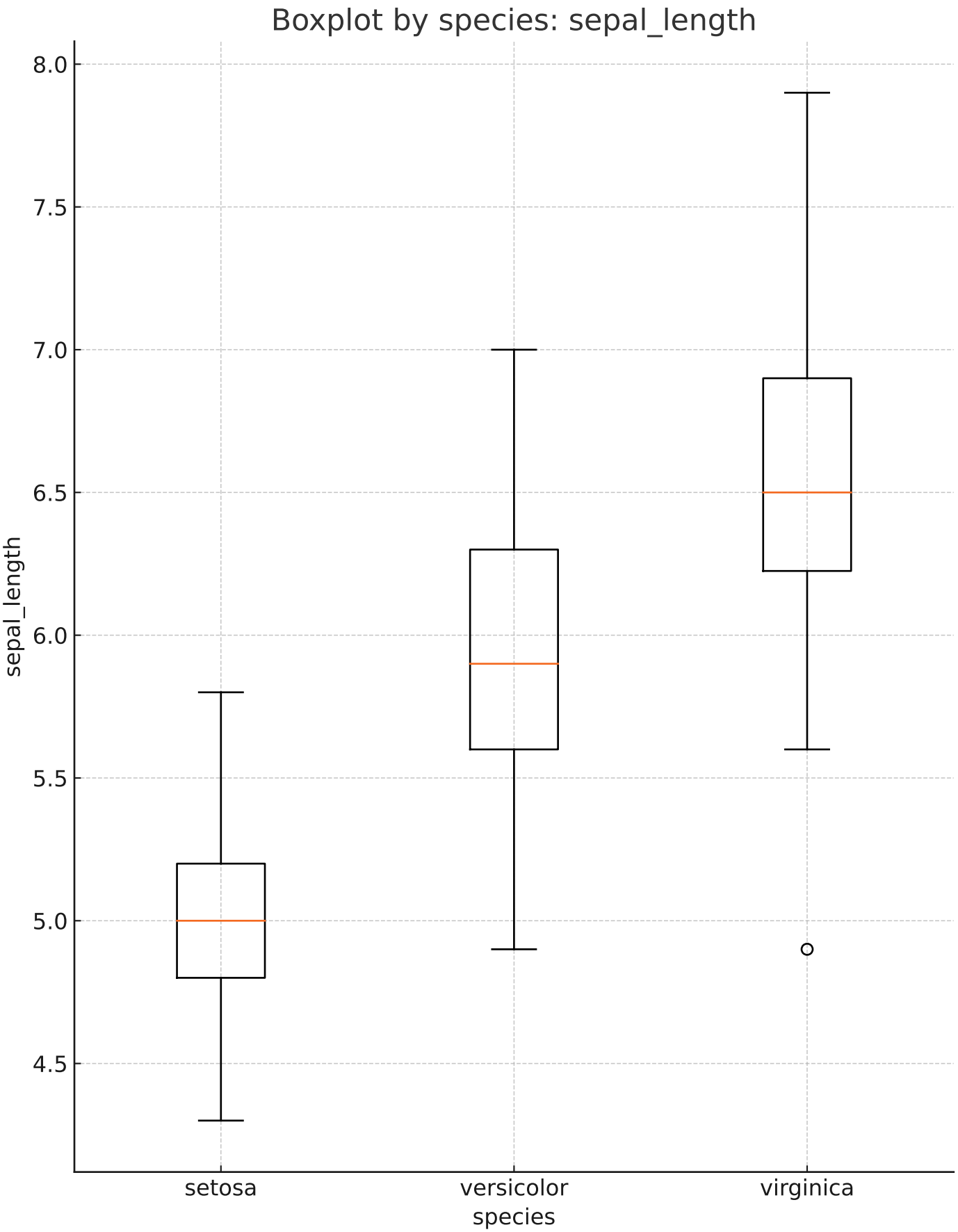


Observations:
Sepal width is slightly left-skewed with a thicker left tail below 3.0.

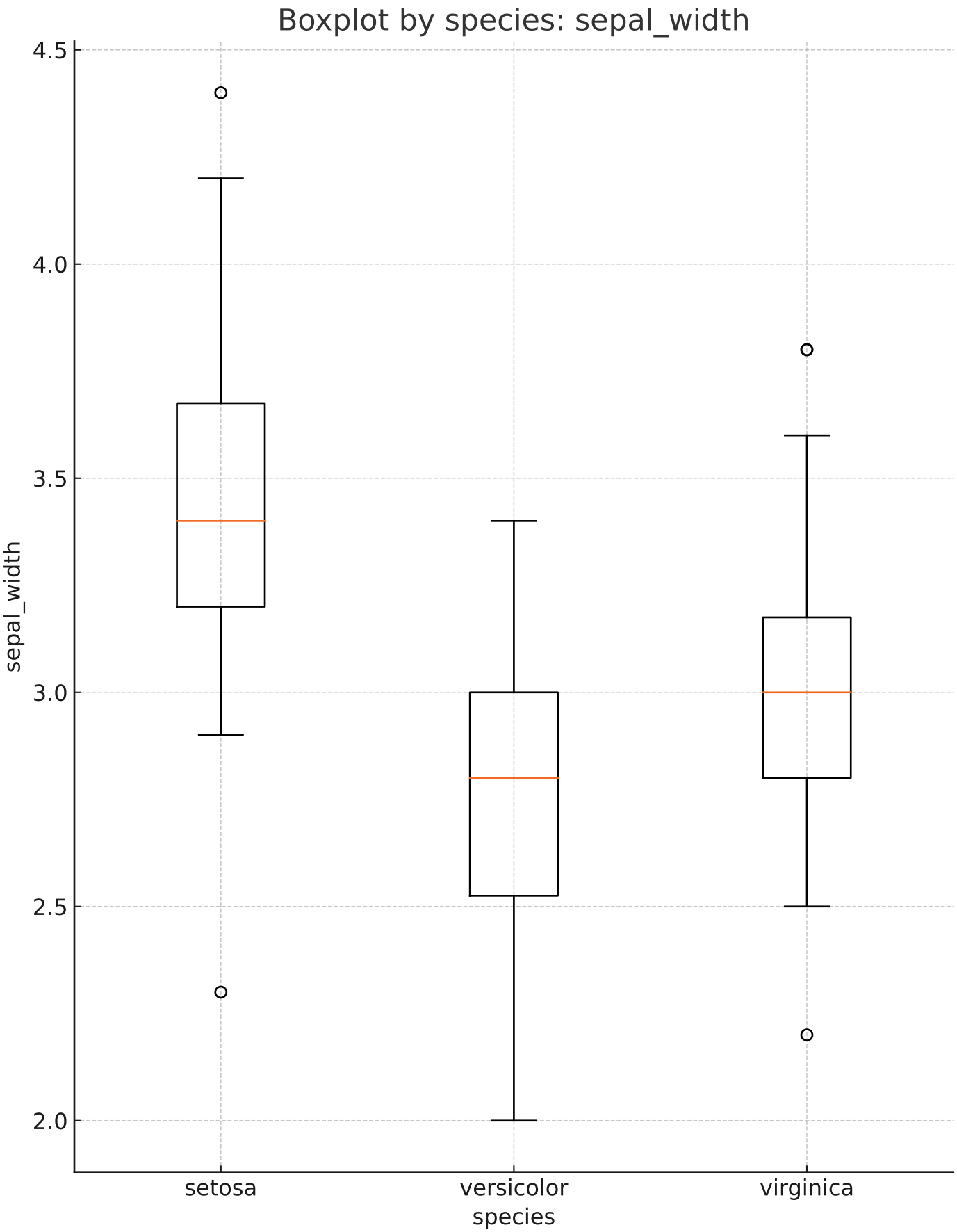
Observations:
Petal length is clearly bimodal; setosa separates from the other two species.



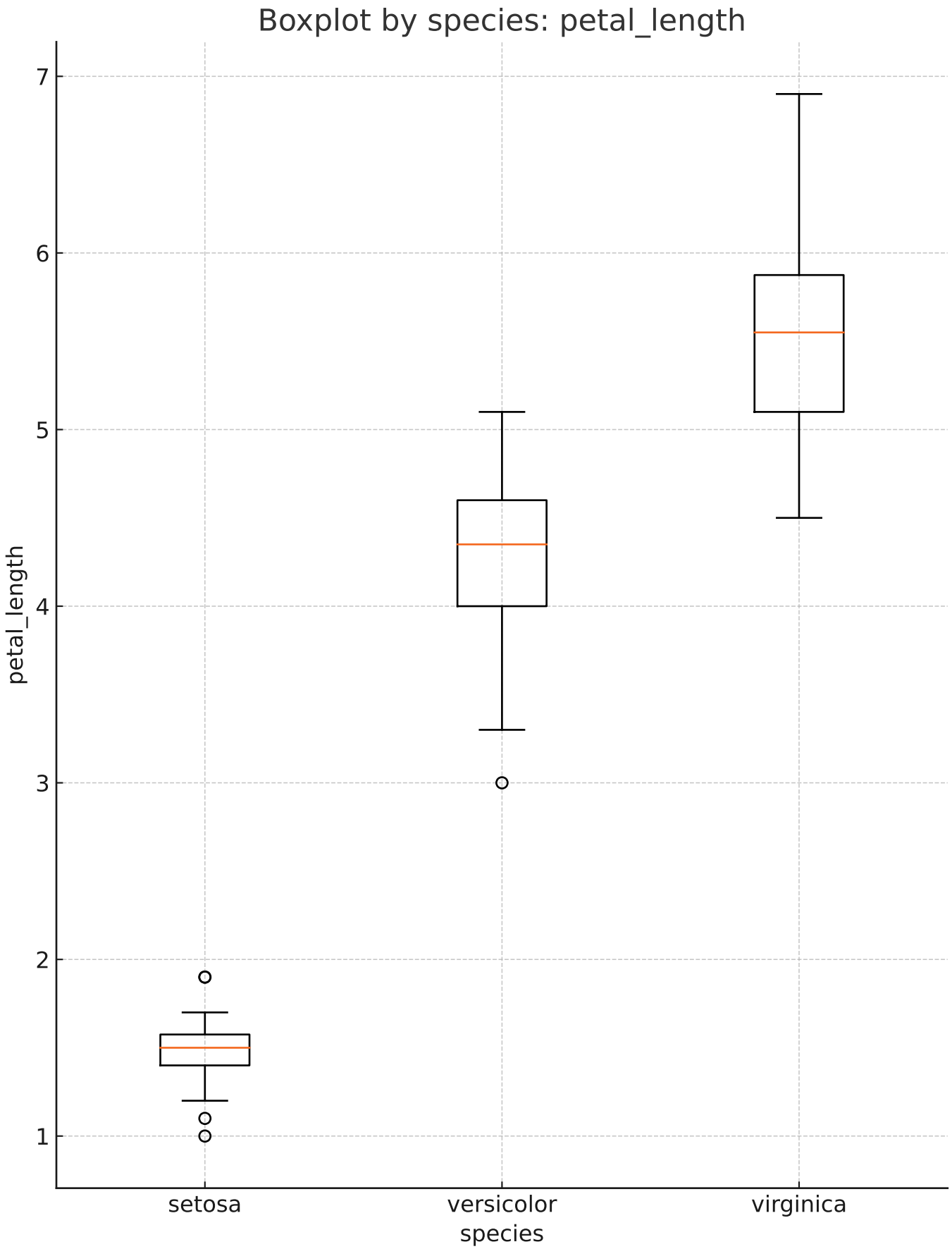
Observations:
Petal width is also bimodal; strong separation for setosa.



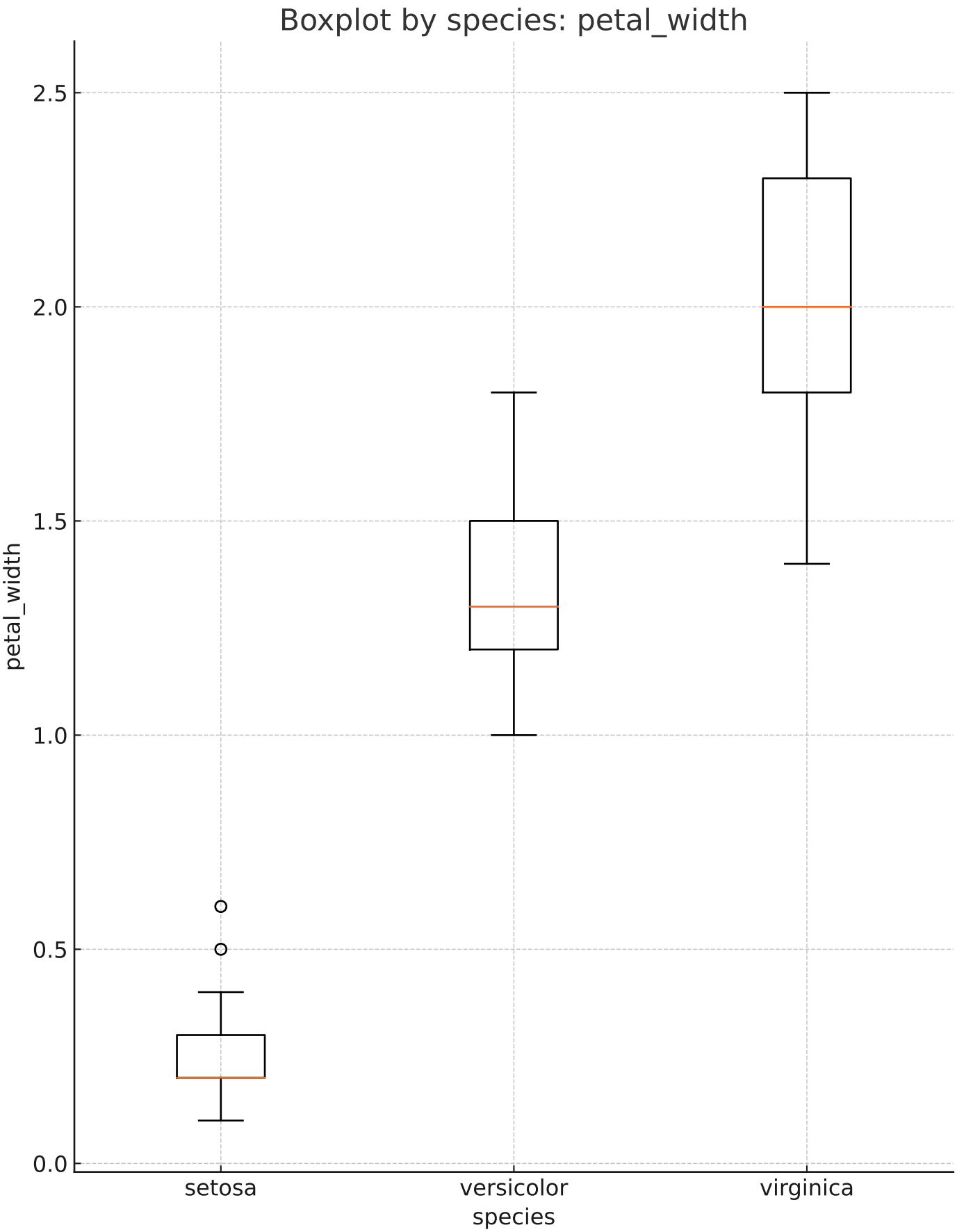
Observations:
Versicolor and virginica have higher median sepal lengths than setosa; some overlap exists.



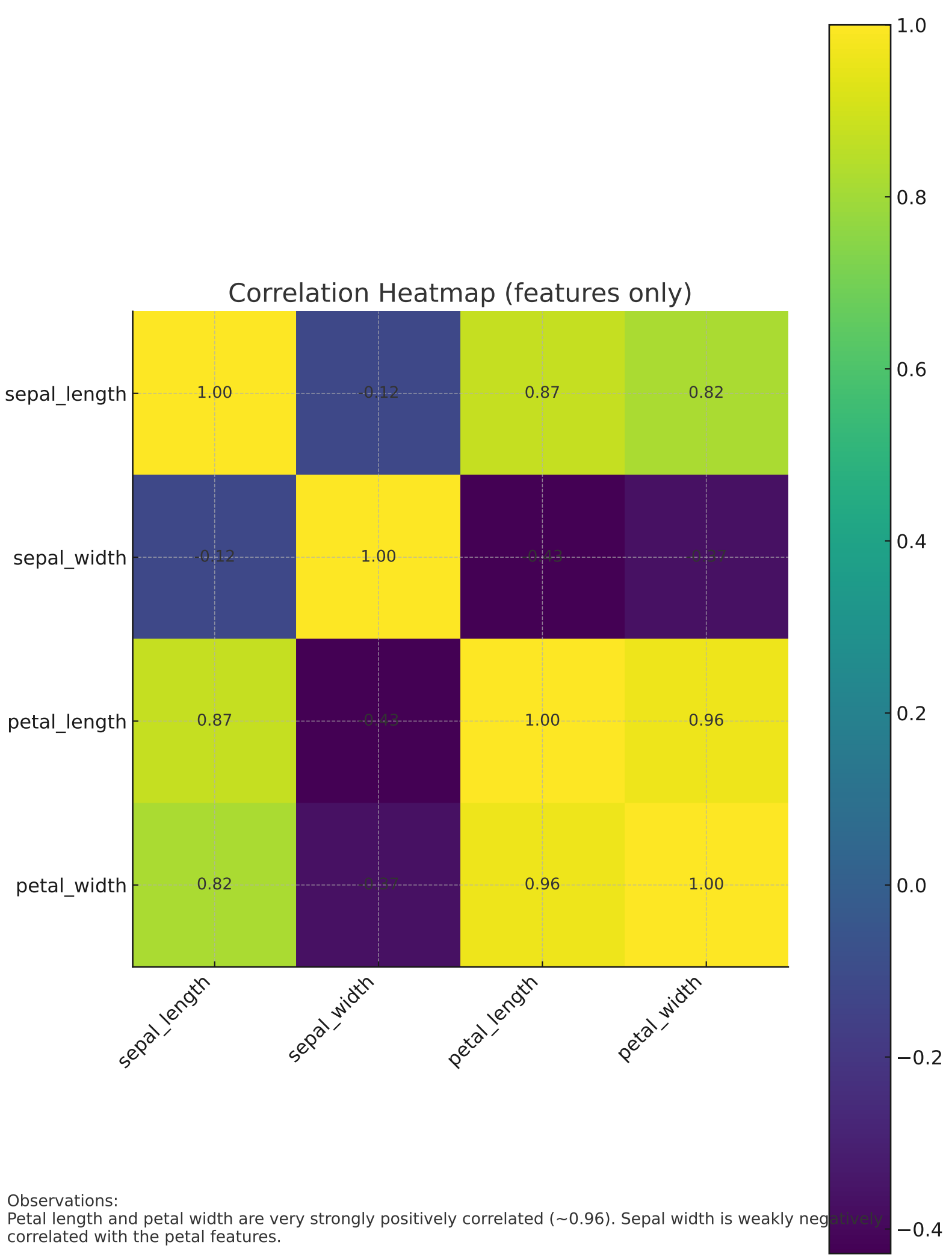
Observations:
Setosa tends to have slightly higher sepal width median compared to the others.



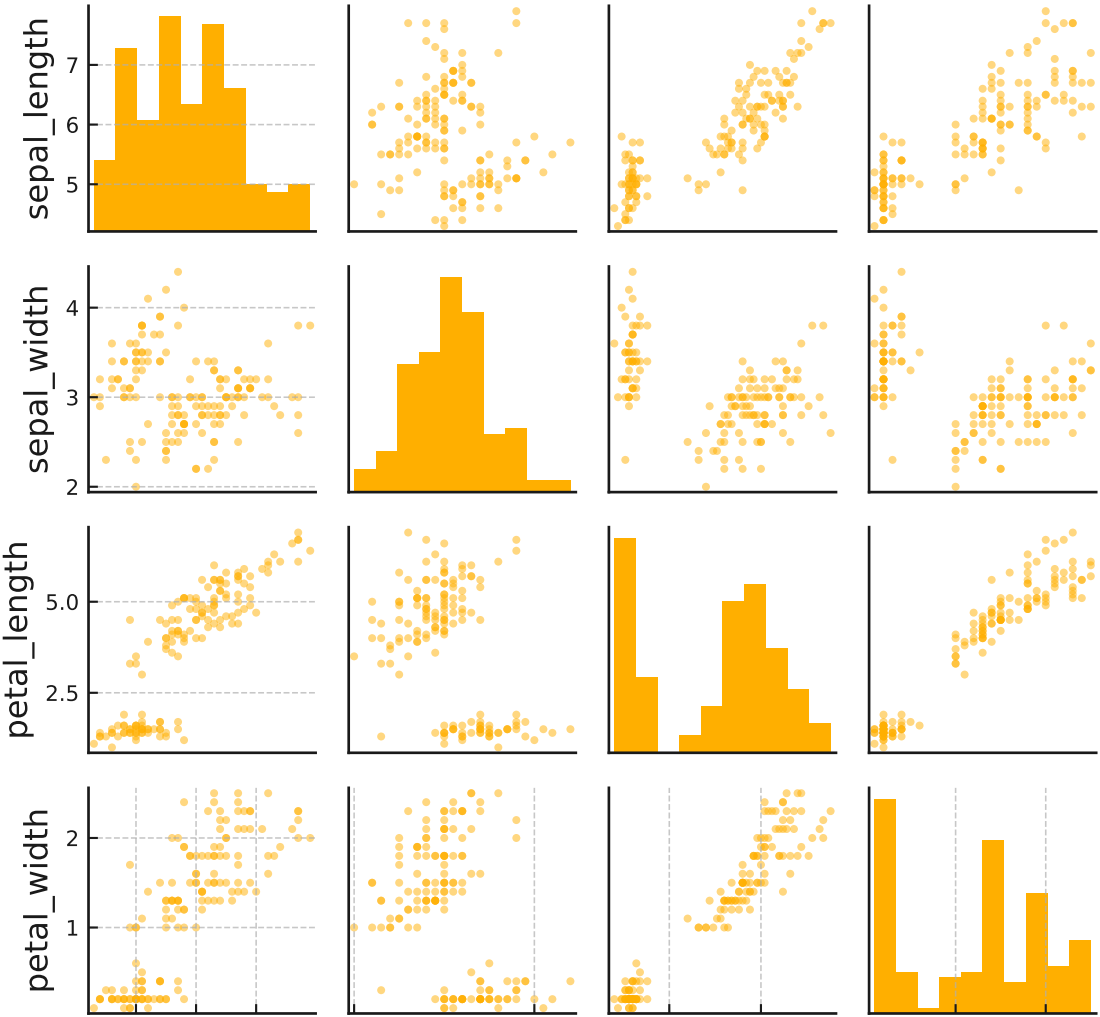
Observations:
Species are well separated by petal length; setosa smallest, virginica largest.



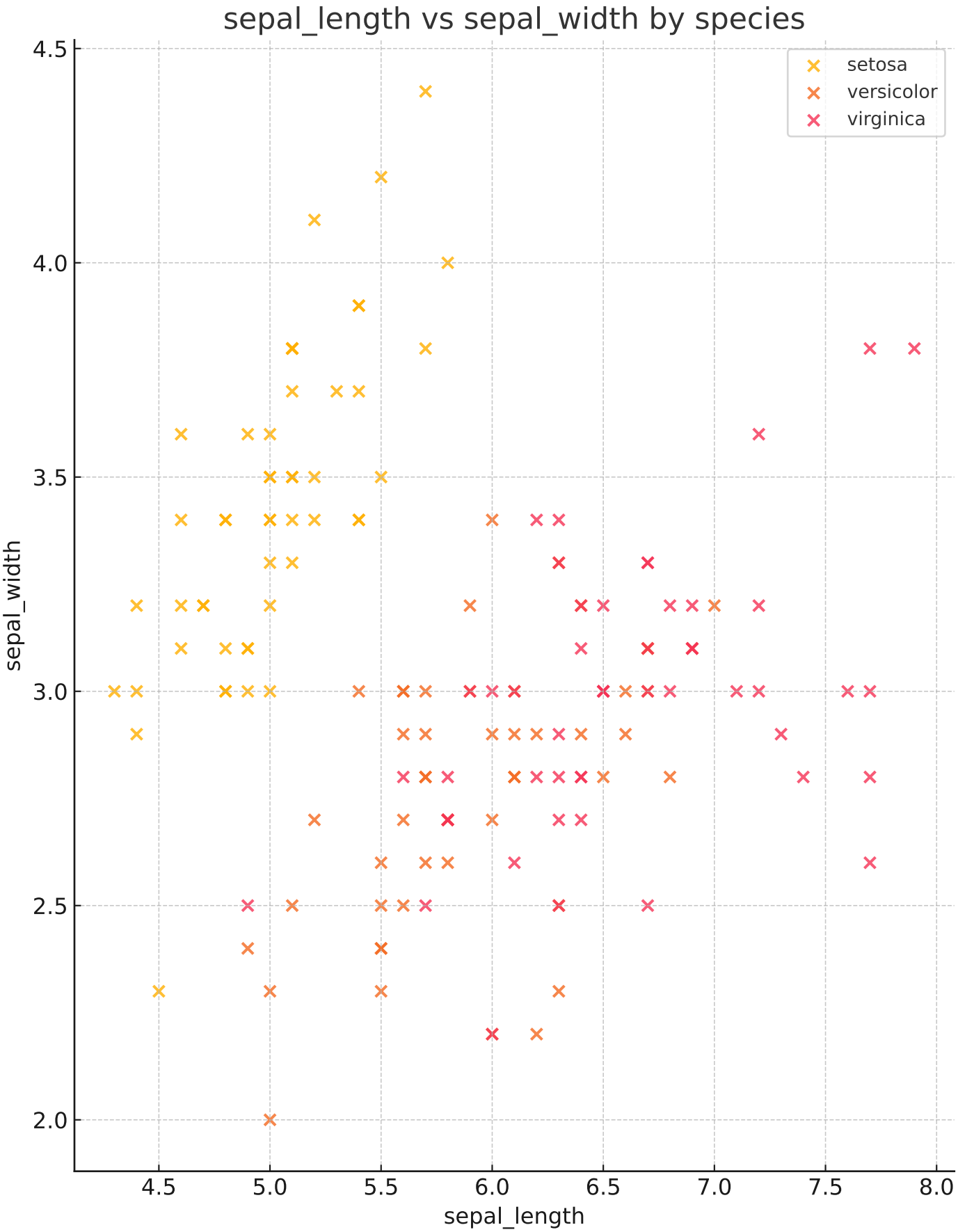
Observations:
Similar clear separation by petal width; potential for strong predictive power.



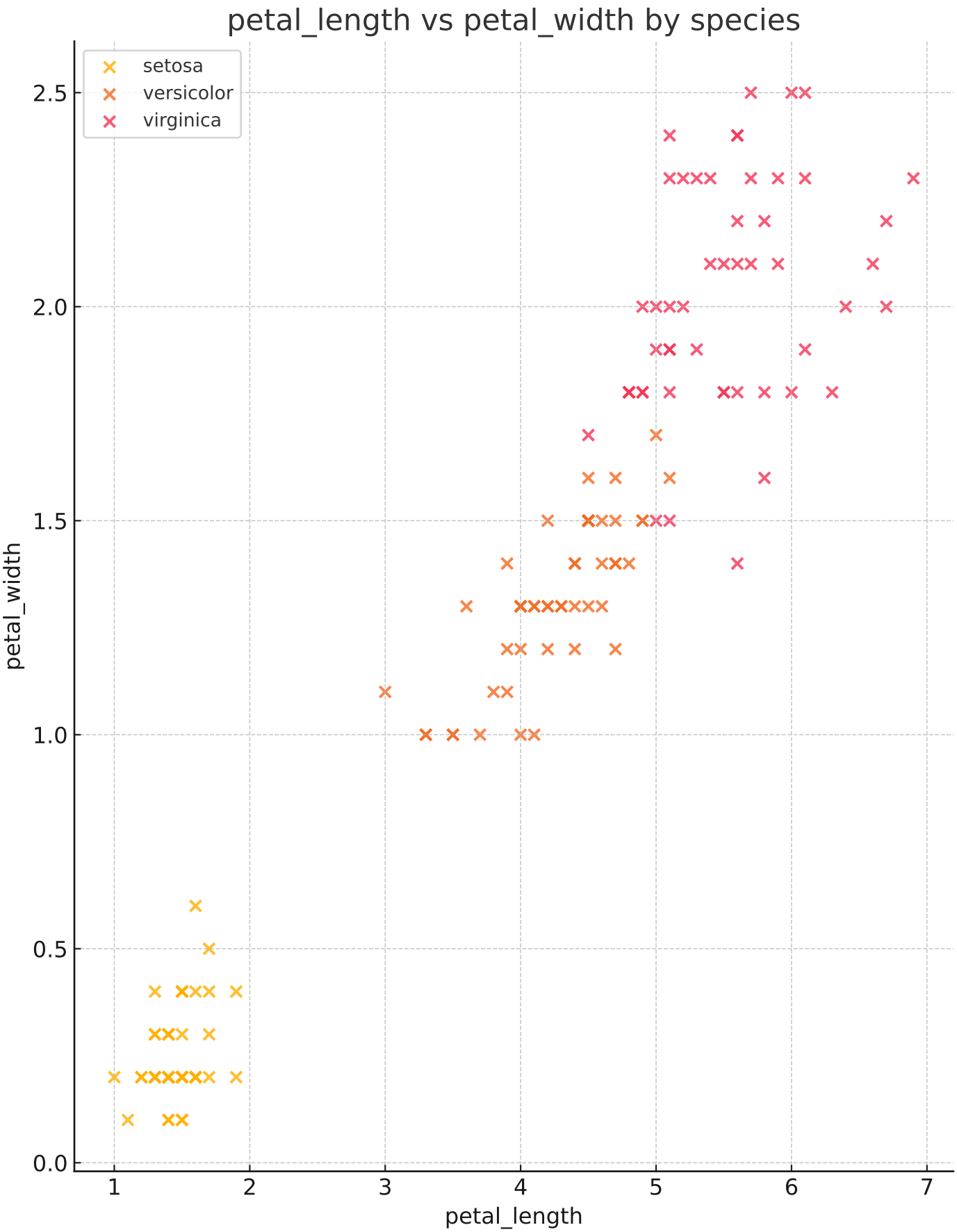
Observations:
Petal length and petal width are very strongly positively correlated (~0.96). Sepal width is weakly negatively correlated with the petal features.



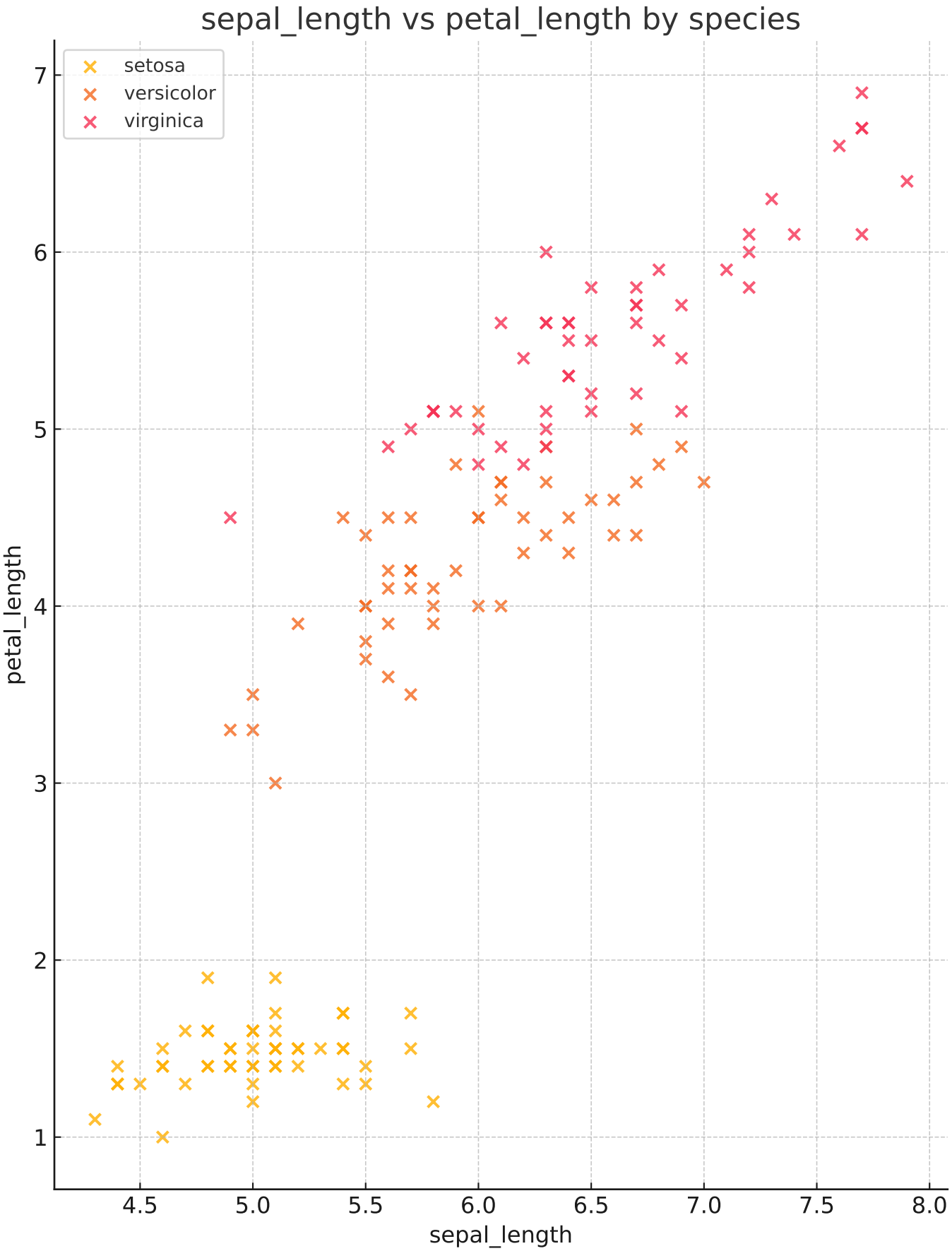
Observations:
Petal dimensions show tight linear relationships; sepal relationships weaker and more dispersed.



Observations:
No strong linear relation overall; setosa shows slightly higher sepal width for given sepal length.



Observations:
Near-linear relationship with distinct clusters per species.



Observations:
Positive trend; species clusters are evident, especially setosa vs others.